

A multi-resolution and multivariate analysis of the dynamic relationships between crude oil and petroleum-product prices

Josué M. Polanco Martínez^{*a,c}, Luis M. Abadie^a, J. Fernández-Macho^{b,c}

^a*Basque Centre for Climate Change - BC3, 48940 Leioa, SPAIN.*

^{*}*Corresponding Author: josue.m.polanco@gmail.com*

^b*Department of Econometrics and Statistics, University of the Basque Country, 48015 Bilbao, SPAIN*

^c*Institute of Public Economic, University of the Basque Country, 48015 Bilbao, SPAIN*

Abstract

This paper proposes the use of a novel multivariate, dynamic approach – wavelet local multiple correlation (WLMC) [1] – to analyse the relationship between oil time series in the time-scale domain. This approach is suitable for use with energy data of any kind that change over time and involve heterogeneous agents who make decisions across different time horizons and operate on different time scales. The study of the links between multivariate oil time series is of great importance in energy research, e.g., it is extremely important for petroleum industry refiners and investors to know the relationships and margins between output prices and crude oil costs. The estimation of wavelet correlations in a multivariate framework between such prices is a suitable way to analyse crude oil and petroleum products as a system. To exemplify the use of WLMC, we analyse the relationships between the prices of seven commodities: West Texas Intermediate crude oil and six distilled products (conventional gasoline, regular gasoline, heating oil, diesel fuel, kerosene and propane) from 10/06/2006 to 17/01/2017. The results reveal that the wavelet correlations are strong throughout the period studied and there is a strong decay in correlation values from 2013 to 2015. The most plausible explanation for this decay is overproduction of tight oil in the U.S. and a slowdown in global demand for oil. Furthermore, our results also reveal that heating oil, diesel and kerosene maximise the multiple correlation with respect to the other oil variables on different scales, indicating that these products are the most dependent variables in the crude-product/price system. WLMC offers new opportunities for applications in energy research and other fields.

Keywords: Oil spot prices, oil refined oil product prices, wavelet correlation, MODWT, Non-stationary time series

1. Introduction

Crude oil is an important, indeed indispensable commodity for the global economic system [2, 3]. It is one of the most major energy sources for current civilization, accounting for around 40% of global energy consumption [4]. It is used in different economic activities and domains including industrial production, transportation, and agriculture,

among others [5]. Crude oil is not a homogeneous commodity and is classified by its density and sulphur content. Taking into account these physical and chemical properties, there are over 160 different internationally traded crude oils [6]. Two of the most representative worldwide crudes are West Texas Intermediate (WTI) and Brent, which are the benchmarks for the US and Europe [3, 6]. Crude oil can be refined into many petroleum products, e.g., regular gasoline, heating oil, diesel fuel, jet fuel and others. These refinement process converts approximately 47% of crude oil barrels into gasoline, 24% into diesel fuel and heating oil, 13% into jet fuel oil (e.g., kerosene), 4% into heavy fuel oil (HFO), 4% into liquefied petroleum gas (LPG; e.g., propane) and 8% into other products (e.g., paraffins or asphalt) [7, 8].

The prices of crude oil and the main petroleum products (gasoline, diesel and heating oil) are the ultimate embodiment of trading results in the global market [9]. Crude oil price fluctuations have significant impacts mainly on economic growth [10], financial markets [5, 11] and national security [4]. Crude oil prices are driven by different market factors such as supply and demand; moreover they are strongly influenced by exogenous factors such as irregular events [12], global economic status [13], speculation [14], and political and social attitudes [15], the effects of which on the crude oil market are not always easy to quantify [9]. On the other hand, the petroleum-products prices are enormously influenced by crude oil prices [16, 17] and by demand from consumers, the strategies of refineries and investors, stocks volumes, seasonality and meteorological conditions (e.g. heating oil) [17, 18]. The prices of the main petroleum products can also influence crude oil prices [17, 19]. This means that there is feedback between them. For this reason, strong correlations between the prices of crude oil and petroleum products are expected, as pointed out previously by several authors [6, 19, 20].

The study of the relationships between the prices of crude oil and refinery products is of great interest to refiners, investors, policymakers and scholars. One of the main reasons for that interest is that it is extremely important to know the differences or margins (or crack spread) between the output prices of distilled products for petroleum industry refiners and crude oil costs. It is therefore fundamental to investigate what dynamic interactions there are between crude and refined product prices. Since the seminal works of Girma and Paulson [21] and Gjølberg and Johnsen [22] published in 1999, a relatively moderate amount of research effort has been done during the last 20 years. For example, [21] and [22] analyse the co-movements between the prices of crude oil and major refined products and investigate the long-term pricing relationship. [21] finds that these commodities are co-integrated for the period 1983–1994, whereas [22] finds that the spread was stationary during 1992–1998. The next relevant paper was published in 2003: [16] apply a multivariate approach in the time domain to study the relationships between the prices of crude oil and several refined products for 1992–2000. They find that crude oil prices are weakly exogenous, that the spread is constant in some but not all relationships, and that the link between these prices implies market integration for these refined products. Shortly afterwards, in 2005 [6] carried out one of the most complete studies about this topic: they investigated ten prices series of crude oil and fourteen prices series of petroleum products from America and Europe from 1999 to

2002 through co-integration and error correction models. They find that the significance of product prices in the long-run relationship is specific to each geographical area, the relevant product mix also depends on the market area and on the characteristics of the crude, and the market price is the driving variable of the crude price also in the short-run, irrespective of the specific area and the quality of the crude.

Despite the importance of the study of the links between the prices of crude oil and its products, the number of papers published from 2005 to the beginning of 2010 is small, especially as regards papers in which several products are analysed, in comparison, for example, to the number of studies published about the link between crude oil spot prices and futures markets [23]. Some exceptions are [24], who studied retail gasoline and crude oil price movements in the US and find that gasoline prices in the long-run are influenced more by the technological changes on the demand side than by crude oil price movements on the supply side. It is worth noting that this finding is in disagreement to the literature that attributes asymmetric price movements to market power refiners. However, we would like to highlight that [24] only analyse the prices of one petroleum product and the paper does not consider crude oil and petroleum products as a system. Other exceptions are [8], [19] or [23], and although some products are analysed in these papers they focus more on crack spreads than on the empirical relationships between the prices of crude oil and petroleum products.

During the last few years, however, there has been a relative increase in the number of publications that treat the prices of crude oil and the main products as a system, and in particular in paper that use a variety of advanced statistical techniques originally developed to analyse complex systems. For instance, [17] uses a copula approach to capture potential nonlinear relations between crude oil and some refined products and finds that lower and upper tail dependences are both positive, indicating that crude oil and refined product markets tend to move together; the same paper also finds asymmetry in tail dependence between crude oil and heating oil/jet fuel returns. [18] uses detrended cross-correlation analysis to investigate the cross-correlations between crude oil and refined product prices for 1991–2013, and finds that cross-correlations are significant and strong, that there is strong multifractality in them, and that long-term cross-correlations are stronger in recent ten years than in previous decades. [25] also investigates the same problem and finds that nonlinear correlations are stronger in the long-term than in the short-term, crude oil and product prices are co-integrated and financial crisis in 2007–2008 caused a structural break in the co-integrating relationship so that refined product prices dominated crude oil prices during the financial crisis. [26] analyses a dynamic conditional correlation between crude oil and fuels prices in a non-linear framework and finds a need for a different model for each pair analysed and the presence of at least one structural break in the conditional volatility and in the correlation between WTI and each product.

Such statistical techniques are suitable way of obtaining knowledge in energy research due the fact that oil/petroleum products prices make up a complex system [10, 11]. This means that there are a large number of variables and factors interacting with one another and belonging to a range of areas that are interconnected, e.g. energy industry,

economics, finance, engineering, technology, environment, etc. [27, 28, 29]. On the other hand, these techniques are specifically designed to tackle many of the characteristics contained in oil market data, such as the following: 1) oil time series change over time, i.e., they are not stationary (in other words their main statistical properties, e.g. their mean or variance, can change over time); 2) these series are not necessarily normally distributed [25, 30]; 3) they can display nonlinear structures [25, 31]; and 4) and they can involve heterogeneous agents who make decisions with different time horizons and operate on different time scales (frequencies) [31, 32]. The wavelet transform (WT) is a suitable statistical technique for handling many of these features contained in oil time series [33, 34].

The wavelet transform is a statistical tool that can handle non-stationary time series and that works in the combined time-and-scale domain (multi-scale analysis). There are essentially two approaches to estimating the WT: the first uses the continuous wavelet transform (CWT) and the second the discrete wavelet transform (DWT) [33, 34]. The CWT enables the spectral features of the time series under analysis and their co-movements to be visualized as a function of both time and scale (frequency) and is adequate for extracting low signal/noise ratio and periodical components. However, it is highly redundant in both time and scale, which is not a desirable feature for correlation decomposition over timescales [1, 34]. By contrast, the DWT is a compact representation of the data, that is, selects a minimal subsample of time-frequency values from the CWT without losing any information contained in the original data, and is particularly useful for noise reduction and data compression [1, 34]. However, in many applications time redundancy is desirable to some extent as long as it allows for data features to be properly aligned and compared across all scales/frequencies. In this sense, the maximal overlap discrete wavelet transform (MODWT) is the most popular wavelet transform as it is redundant in the time dimension but non-redundant in the scale/frequency dimension [1, 33]. Here, we seek to study the relationship between non-stationary oil time series by means of the MODWT, so we are interested in the following methods: wavelet correlation [33, 35], wavelet multiple correlation [36, 37], rolling-window wavelet correlation [38, 39] and, more recently, wavelet local multiple correlation (WLMC) [40].

WLMC measures a non-stationary time-evolving correlation structure at different scales within a multivariate set of data. The concept of wavelet local multiple correlation is used to produce a single set of multi-scale correlations over time, in contrast with the large number of wavelet correlation maps that need to be compared when standard pairwise wavelet correlations with rolling windows are used [1, 38]. After the WLMC is applied to multivariate data sets, two kinds of useful outputs are obtained: 1) a heat map where the wavelet local multiple correlations are drawn; and 2) a set of plots (which depends on the number of wavelet scales obtained after apply the wavelet transform is applied to the data under study) where the variable that maximises the multiple wavelet local correlation in each wavelet scale is highlighted. The other advantage of WLMC is that the computational implementation, in R language, is freely available as an R package on CRAN [1], which enables any researcher to reproduce our results presented in this paper and to apply this innovative methodology to their own problems and data

sets.

This novel tool is useful for analysing the link between prices of crude oil and petroleum products, which is clearly justified because the links between crude oil and its derivatives can change over time [26]. Furthermore, some recent studies have shown that the link between oil price and markets of other kinds, e.g. the exchange rate, is time-frequency dependent and bidirectional at the same time [32, 41]. In line with these studies and this argument, our paper proposes for the first time the use of wavelet local multiple correlation [40] to analyse the interrelationships among non-stationary multivariate oil time series such as the prices that come from the complex system formed by crude oil and petroleum products.

The objective of this paper is to use a novel multivariate, dynamic approach – wavelet local multiple correlation (WLMC) [1] – to analyse the link between oil time series in the time-scale domain. As a case study, we analyse dynamically the links between the prices of seven commodities: crude oil prices (West Texas Intermediate (WTI) spot prices, one of the world’s main crude oil benchmarks) and the prices of six refined products (conventional gasoline, RBOB regular gasoline, heating oil, Ultra-Low-Sulphur diesel fuel, kerosene and propane) using daily market prices from 14/06/2006 to 17/01/2017. We are specially interested in the role of the subprime financial crisis and, particularly, the growing tight oil production. As an exploratory analysis, we start by computing the correlation matrix using Spearman’s rank correlation coefficients for the variables under study applied to whole period (2006–2017) and for four sub-periods (2006–2008, 2008–2009, 2009–2014, and 2015–2017) chosen in such a way that they cover certain financial/energy events of interest. We chose the Spearman method because it does not assume normal distribution for the variables analysed and oil time series are not necessarily normally distributed [25, 31], but other estimators of correlation can be used. The purpose is to show that the correlations between oil time series change over time, suggesting the need for a more sophisticated statistical technique to analyse data from complex systems, such as WLMC. Finally, we applied the recently developed wavelet statistical tool named wavelet local multiple correlation [1, 40] to our case study.

To the best of our knowledge, this innovative statistical tool has never before been used to study oil markets (crude and products prices) in a multivariate, multi-scale context. The results obtained through WLMC could be used to draw up a more efficient period-oriented strategy when the market participants have different temporal horizons (investors, speculators, hedge funds, refineries, etc.) and also for energy policy. In addition, the proposed methodology can be used to improve hedging performance depending on the temporal horizon.

The rest of this article is organized as follows. Section 2 presents a brief literature review about wavelet correlation and oil time series. Section 3 describes the methodology employed. Section 4 presents the case study. Finally, Section 5 concludes.

2. Wavelet correlation between oil time series: literature review

The use of wavelet correlation to analyse relationships between energy time series is relatively new, although a quite number of papers have been published in the last

five years. These can be classified in three large groups: the first group explores the links between crude oil and various stock markets, exchange rates and other markets commodities (such as precious metals, like gold). Among these studies, the following stand out: [42] uses wavelet multi-resolution analysis (wavelet correlation and cross-correlation through MODWT) to examine the relationship between oil and stock markets in Europe and the USA at the aggregate and sectoral levels for the period from June 2000 to July 2011. [43] uses a combination of the Haar (discrete) wavelet transform and vector autoregressive models to examine the asymmetric effect of oil prices on stock market but taking into account various time horizons. More recently [44] reports an interesting study into the role of several oil prices on stock markets, providing a novel approach combining wavelet coherence and a network model. Moreover, [45] uses the Haar (discrete) wavelet transform, wavelet coherence and multiple wavelet coherence through the CWT to analyse daily Brent oil prices, London gold fixing price and the Shanghai Composite index from January 1991 to September 2014. It is noteworthy that this publication uses multiple wavelet coherence developed by [46], which is able to analyze multivariate time series. Last but not least important, [32] uses wavelet coherence (via the CWT) and multi-horizon Granger causality testing between the returns on oil prices and the India-US exchange rate over the time interval 1980–2016. One of the most interesting methodological advances in this paper is the use of Granger-causality tests proposed by [47] and implemented within the wavelet coherence, which avoid applying the Granger test to each wavelet decomposition coefficients of the time series under study.

The second group contains an analysis of the links between future oil prices. This group includes papers such as [48], where the authors study the relationships between the daily closing prices of the near month futures contracts for light sweet crude and natural gas for the period 1990–2007. [48] uses the CWT to obtain the wavelet coherence and cross bicoherence for the energy data analysed. Wavelet bicoherence was originally developed by [49] to analyse turbulence (physics of fluids) but it has not been widely used in the analysis of energy time series. [50] studies the relationships between daily prices of NYMEX light sweet crude oil with one-month maturity contracts and daily spot prices of WTI oil in the interval 2003–2011 by means of linear and non-linear causality tests applied to the wavelet decomposition coefficients of these two data sets. More recently, [31] analyses daily spot prices of crude oil (WTI) and long-term futures (calculated as the expected value from a stochastic model calibrated with the futures quotes of each sample day) for the period 2006–2016 through the wavelet cross-correlation via the MODWT and a non-linear causality test applied to the wavelet decomposition of the WTI and futures.

The third and last group comprises publications that analyse the relationships between several crude oil prices. For instance, [51] propose an original approach involving the combination of the wavelet correlation via MODWT and grey correlation (originally proposed by [52] and able to reveal the dynamic relationship between non-linear time series), a multivariate dynamic framework for analysing the correlation between four crude oil time series (Brent, Dubai, Minas and Daqing). Recently, [3] also propose a novel approach based on wavelet analysis through MODWT and complex networks to learn about the evolution of many world crude oil markets (26 weekly crude oil groups) for the period

1999–2011. Following the same philosophy, i.e. that correlation between two or more energy time series is dynamic (correlation can change over time and scale or frequency), here we contribute to the literature by proposing wavelet local multiple correlation [40], a novel multi-scale, multivariate method for analysing the links between non-stationary energy time series in both time and scale/frequency at once. On the other hand, despite the large body of literature analyzing the relationship between various kinds of crude oil prices, no papers have been published that analyse the relationships between the prices of crude oil and products through a multivariate framework such as WLMC.

3. The wavelet local multiple correlation

Wavelet local multiple correlation (WLMC) measures a non-stationary time-evolving correlation structure at different scales among a multivariate set of data [1, 40]. Alternatively, combining standard bivariate wavelet correlation analysis with rectangular rolling time windows require calculating, plotting and comparing a large number of wavelet correlations that would now require an additional time dimension [53, 54, 55]. More specifically, with n time series, a total of $n(n-1)/2$ wavelet correlation heat maps would be required, each of dimension $J \times T$, where $J = \lfloor \log_2(T/(L-1) + 1) \rfloor$ is the order of the wavelet transform, L is the wavelet filter length and T is the time series length. This soon becomes exhausting and confusing.¹ By contrast, the WLMC method based on multiple regression consists of a single set of multi-scale relationships which can be expressed in single sets of scale-time local correlation heat maps. This is not only easier to handle and interpret but it may also provide a better insight into the overall statistical relationships and co-movement dynamics within the multivariate time series under scrutiny. Moreover, some consideration as to appropriateness of the usual rectangular rolling window needs also be given and, the WLMC accordingly offers a choice of six different windows with different spectral properties.

Wavelet analysis, like Fourier analysis in the frequency domain, decomposes data into components at different frequencies, the difference being in that it offers flexible resolution depending on the time scale of each frequency component. Because of this flexibility, in the last 30 years it has become a popular tool in a wide range of scientific fields such as acoustics, climatology, imaging, signal processing, seismology or even, in recent times, economics, finance and energy research. In fact, it is widely used in many applications where it is felt that data may have different influences and behavior depending on the time horizon.

For a multivariate time series or signal $x(t)$ the wavelet transform is defined similarly to the short-time Fourier transform that seeks to reach a compromise between time and frequency by focusing on windowed portions of the time series. Wavelets add further flexibility by replacing the time window with scaled-and-shifted versions of a mother

¹For example, even the moderately large number of series required in the empirical analysis of Euro-zone stock markets [1, 56] would have to handle a total of 55 wavelet correlation heat maps of dimension $[9 \times 4542]$, which would render pairwise multi-scale comparisons pointless in practice.

wavelet $\psi(u)$ with compact support. Thus, the continuous wavelet transform (CWT) can be written as

$$w_i(s, \tau) = \int_{-\infty}^{\infty} x_i(t) \psi_{s\tau}^*(t) dt, \quad i = 1 \dots n, \quad \text{with} \quad \psi_{s\tau}(t) = (1/\sqrt{s}) \psi((t - \tau)/s), \quad (1)$$

where the mother wavelet $\psi(u)$ satisfies certain requirements such as zero integral and unit energy. This ensures that an admissible mother wavelet $\psi(u)$ is well localised in both time and frequency and behaves like a brief oscillation or wave along the time axis (hence the name “wavelet”).

As such the CWT in 1 is highly redundant, since it depends on two parameters—scale and time—while the original data depend on time only. Therefore, it suffices to observe a discrete number of scales with a varying but also discrete number of coefficients $w_i(s, \tau)$ at each scale to contain, and eventually recover, all the information in $x_i(t)$. Thus, the discrete wavelet transform (DWT) is just a critical sampling of the CWT which selects a minimal sub-sample of time-frequency values from the CWT without losing any of the information present in the original data. This is an important feature for some engineering applications such as signal and image compression, but for other applications, e.g. with economic or financial data, time redundancy is desirable to some extent as long as it allows for data features to be properly aligned and compared across all scales/frequencies. For such cases, the maximal overlap discrete wavelet transform (MODWT) is the most popular wavelet transform as it is still non-redundant in the scale/frequency dimension but is redundant in the time dimension. It is obtained as a sampling of the CWT with $s = 2^{-j}$, where $j = 1, 2, \dots$ is the wavelet scale level.

Furthermore, since each $x_i(t)$ is a discrete time series of length T for most applications, it can be shown that each $T \times n$ matrix of wavelet coefficients $W(j)$ can be obtained from

$$W(j) = \mathcal{W}_j X, \quad j = 1 \dots J, \quad (2)$$

where $X = [x(1), \dots, x(T)]^\top$, $J = \log_2(T)$ is the maximum scale level and the rows of matrix \mathcal{W}_j are made up of circularly shifted versions of the vector of T wavelet filter values $\psi_{j\tau}$ from 1 for a given discrete wavelet filter such as any member of the Daubechies family [57]. It can also be noted that at each level j these wavelet coefficients are associated with changes in the effective scale of length $\lambda_j = 2^{j-1}$, which roughly corresponds to periods in the range of $[2^j, 2^{j+1}]$ time units, while the “smooth” remainder is contained in one scaling coefficient [58]. Apart from its time redundancy, MODWT is known to have several major advantages, including energy preservation, which is particularly important in many applications (57; 59, p.135). The MODWT wavelet coefficients thus obtained from the basis for the wavelet multiple regression statistics calculated by the WLMC.

Consider $W(j) = [w_1(j), w_2(j), \dots, w_n(j)]$ in 2, i.e. the matrix of scale λ_j wavelet coefficients obtained by applying MODWT to each series in a realization of a multivariate stochastic process, and also define, $W_i(j)$ at each scale λ_j as the $T \times n$ matrix obtained by replacing the transform w_i in $W(j)$ by a column of ones. The idea is to obtain a linear

function at each scale λ_j that minimises a weighted sum of squared errors for a fixed $\tau \in [1, \dots, T]$

$$S_{j\tau} = u_{j\tau}^\top \Theta_\tau u_{j\tau} \quad \text{with} \quad u_{j\tau} = W_i(j) \beta_{j\tau} - w_i(j), \quad \forall j, \quad (3)$$

where Θ_τ is the $T \times T$ diagonal matrix whose diagonal elements are the values $\theta(t - \tau)$ for a given moving average weight function or window $\theta(d)$ that depends on the time lag between observed MOWDT coefficients $w_i(j, t)$ and $w_i(j, \tau)$.

Letting τ move over time, T local linear regression fits are obtained at each scale λ_j , each with its corresponding residual weighted sum of squares

$$\hat{S}_{j\tau} = \hat{u}_{j\tau}^\top \Theta_\tau \hat{u}_{j\tau}, \quad \hat{u}_{j\tau} = W_i(j) \hat{\beta}_{j\tau} - w_i(j), \quad \forall j, \tau. \quad (4)$$

where $\hat{\beta}_{j\tau}$ are the least-squares estimates of the unknown parameters $\beta_{j\tau}$ in 3. These residuals can then be used to calculate J series of local coefficients of determination

$$R_{j\tau}^2 = 1 - \hat{S}_{j\tau} / TS_{j\tau}, \quad \tau = 1 \dots T, \quad j = 1 \dots J, \quad (5)$$

where $TS_{j\tau}$ is the total weighted sum of squares at time τ and scale λ_j that measure how good the least-squares fit in 3 at each time τ and scale λ_j .

Finally, the WLMC is obtained as a collection of local multiple correlation coefficients $\varphi_X(\lambda_j, \tau)$ calculated as the square roots of each of the coefficients of determination $R_{j\tau}^2$ corresponding to the linear combination of weighted variables $\Theta_\tau^\frac{1}{2} w_{ij}$, $i=1 \dots n$, where that local coefficient of determination is a maximum. To estimate the WLMC we use the *wavemulcor* R package, which is freely available from the CRAN repository [40].

4. Case study: the crude-product/price system

4.1. Data description

The data used in this study comprise daily market quotes for West Texas Intermediate (WTI), conventional gasoline (N.Y.), RBOB regular gasoline, heating oil, Ultra-Low-Sulphur diesel Fuel (N.Y.), kerosene and propane. All data sets cover the period from 16/06/2006 to 17/01/2017 (2668 trading days). Data were obtained from the U.S. Energy Information Administration (EIA)². Figure 1 shows how prices change over time, with prices per gallon are converted to dollars per barrel using 1 barrel = 42 U.S. gallons, based on U.S. production (2014 Energy Information Administration (EIA)). It is clear that crude oil and refined product prices have roughly similar dynamics over time, with the exception of propane. This can be explained straightforwardly due to the fact that these refined products are actually “fractions” of WTI crude oil. It can be clearly observed that propane only accounts for 4% of the crude oil refined. Thus, propane prices can to some extent be expected not to be so similar to the WTI and to the other products. On the other hand, propane is affected by different fundamentals and drivers in the U.S, e.g.

²https://www.eia.gov/dnav/pet/pet_pri_spt_s1_d.htm

it is highly seasonal and is sensitive to meteorological conditions and catastrophic events, while the WTI is not influenced by seasonal patterns. Moreover, propane production comes largely from two industrial sectors: crude oil refining and natural gas processing and extraction. Production from gas plants is currently greater than from refineries. Therefore, propane prices are also affected by natural gas processing [60, 61]. In addition, [26] confirm the presence of structural breaks in the volatility series and in the correlations between WTI returns and the returns on gasoline, kerosene jet fuel, diesel, heating oil, propane and natural gas. Their results show that a shock in the price of crude oil takes longer to be transferred to diesel and propane than to the other products.

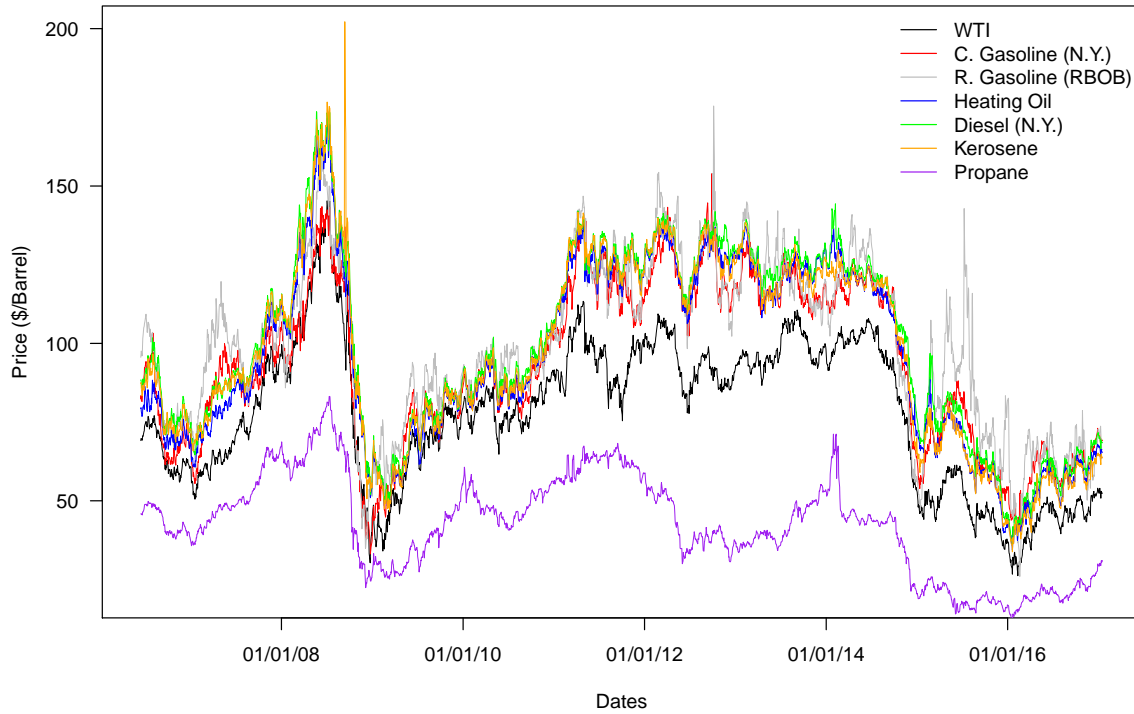


Figure 1: WTI crude oil prices and prices of six refined products for 14/06/2006 – 17/01/2017 (daily data). Source: own work using data from EIA.

Basic descriptive statistics for crude oil and refined product prices for the period under study (14/06/2006-16/02/2017) and a Jarque-Bera test for normality as implemented in the R package Tseries [62] are presented in Table 1. There are three distinct groups of mean values: 1) WTI; 2) the first five refined products (with the highest mean values); and 3) propane (with the lowest mean values). The most volatile variables according to their standard deviations, are the refined products (kerosene, diesel and heating oil at the top and RBOB gasoline and conventional gasoline lower down) followed by WTI. By contrast, propane shows the lowest volatility (this can also be seen in Figure 1). Skewness

Statistics	WTI	C. G (N.Y.)	R.G.(RBOB)	Heating Oil	Diesel (N.Y.)	Kerosene	Propane
Prices							
Mean	77.80	93.18	99.43	95.67	99.68	97.04	42.36
Max.	145.31	153.93	175.43	171.49	173.63	202.19	83.16
Min.	26.19	33.10	25.96	34.31	36.20	33.77	12.43
Std. Dev.	23.31	25.86	26.12	29.27	29.43	30.09	15.36
Skewness	-0.08	-0.11	-0.14	0.02	0.00	0.02	0.02
Kurtosis	-0.73	-1.13	-0.80	-1.13	-1.06	-0.96	-0.72
Jarque-Bera	60.83	148.39	79.85	141.17	123.44	102.29	57.54
p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 1: Descriptive statistic for daily prices for the data under study for 16/6/2006 – 16/02/2017. Bold numbers indicate p -values ≤ 0.01 .

(a measure of asymmetry or more precisely the lack of symmetry) shows that all the variables except diesel show a (left/right skewed) asymmetric probability distribution. Additionally, the kurtosis values show that none of the energy time series under study has a value close to 3 (the theoretical value for a Gaussian probability distribution) indicating that none of the probability distributions of these time series appears to be normally distributed. To confirm this finding, we performed the Jarque–Bera test of the null hypothesis that the respective probability distribution of the prices were Gaussian (chi-square with $df = 2$). The p -values reported led us to reject the null hypothesis in all cases. We observe that the lack of normality of these prices is consistent with the well-known “stylised facts” of the financial time series, as pointed out in previous studies [30, 63].

4.2. Preliminary analysis through Spearman correlation

As an exploratory analysis, we started our analysis by computing the correlation matrix using Spearman’s rank correlation coefficients for the variables under study applied to the whole period (14/06/2006–17/01/2017). To deal with potential drawbacks occasioned by an overall measure of association, and above all to find out whether correlations between the variables under analysis change over time, we split the period 2006–2017 into four sub-periods. The first runs from 16/06/2006 to 31/12/2008 and covers the onset of the global financial crisis; the second runs from 06/01/2008 to 31/05/2009 and its main feature is that it also covers the global financial crisis, with low crude oil prices (this is one of the time intervals where crude oil prices are lowest); the third sub-period runs from 01/01/2009 to 31/12/2014 and covers the period with the lowest degree of volatility; and the last runs from 01/01/2015 and 16/02/2017 and its main characteristic is the presence of tight oil production and low crude oil prices (as in the second sub-period). Since crude oil and products prices are not necessarily stationary and normally distributed (see Table 1 and Section 4.1) [25, 64], the conventional Pearson correlation is not the most suitable estimator for measuring the potential relationship between crude and product prices, as it may lead to a biased estimation [25, 65]. We therefore used the Spearman

rank-order correlation, which does not demand variables with normal distribution, is based on the ranked values for each variable rather than the raw data and is known to be much more robust (than the Pearson correlation), which means monotonic relationships (for instance, non-linear associations). Spearman’s rank correlation coefficient can be estimated as follows [66]:

$$r_s = \frac{1}{n} \sum_{i=1}^n \left(\frac{R(i) - \bar{R}}{S_{n,R}} \right) \cdot \left(\frac{S(i) - \bar{S}}{S_{n,S}} \right), \quad (6)$$

where \bar{R} and \bar{S} are the sample means and $S_{n,R}$ and $S_{n,S}$ are the sample standard deviations calculated with the denominator n (sample size). The rank correlation (r_s) measures the degree of the monotonic relationship between two time series and takes values from -1 to 1 [66].

We would like to remark that the aim of estimating these correlations in different time intervals is to show that the correlations between oil time series can change over time and that standard correlation techniques are only useful for exploratory analysis: more sophisticated techniques must be used to take into account relationships between two or more variables that change over time. We chose Spearman, but other estimators of correlation can be used.

4.3. Correlation in the time domain

Figure 2a provides the Spearman correlation values between all the price time series and the heat maps of those correlation values for the overall period (2006–2017). The first result that shows up is that, in general, the correlation between all the variables is strong (with values from 0.89 to 0.99), except for propane, which shows the weakest correlations (with values from 0.61 to 0.75). This “weak” correlation between propane and the other variables is expected, as mentioned previously. In addition to other reasons already set out (Sec. 4.1) to explain why propane’s prices do not fluctuate in a similar manner to the other variables analysed, the following should also be considered. Crude oil and propane could show extreme swings in supply and demand, although this does not take place at the same time or with the same intensity. This means that the price of propane, which is a crude oil product, can change over time, but not necessarily in the same way as that of crude oil. Propane is widely used as an input in the petrochemical industry and when its price is relatively high in comparison to that of crude oil demand drops significantly, so propane is replaced in the industry by other petroleum products. Other factors that affects the propane market are the growth in domestic propane supply together with changes in the conversion of fuel-oil furnaces to propane, new building construction (improvements in energy efficiency can reduce propane demand) and competition with natural gas and electricity prices in domestic markets (lower crude oil prices makes propane more competitive).

The correlation is also strong in the relationship between the prices of crude and products, ranging from 0.89 (for RBOB gasoline) to 0.95 (for the other four products not counting propane) (Figure 2a). This strong correlation can be explained roughly by two information transmission mechanisms. First, crude oil is the major input of

the refining process, which means that in crude oil prices increase refining costs and thus lead to higher refined product prices. Second, if greater demand increases refined product prices, refiners are willing to produce more to meet that demand [17, 18, 25]. Usually an increase in demand for refined products results in an increase in crude oil prices. The “low” correlation can be explained as follows (in addition to the points mentioned above; see Sec. 4.1): it is well known that lower crude oil prices do not benefit propane markets; indeed they are expected to have a greater impact on the prices of gasoline, diesel fuel and distillate fuel oil [60]. The second result is that the highest correlations (values between 0.91 and 0.99) are between products (excluding propane) rather than with crude oil prices, which seems to be somewhat contradictory given that crude oil is the input for the refined products. However, this result is expected because the dynamic behaviour of product prices is fairly similar from one to the other (except for propane) (see Figure 1). The third noteworthy result is that the highest correlations (with values around of 0.99) are between heating oil and diesel, followed by diesel/kerosene and heating oil/kerosene. These strong relationships can be explained due to the fact that heating oil is very similar to diesel and is comprised of a mixture of petroleum-derived hydrocarbons. Although heating oil and crude oil follow different seasonal trends their prices are still directly related. As demand for crude oil grows, the price of heating oil increases [67]. Nevertheless, these results must be regarded with caution since the Spearman correlation is an overall measure of association and the financial time series could change their statistical properties in different subintervals (e.g. during periods with high volatility) [25, 38].

The first conspicuous finding, beyond the fact that correlation for the overall period is high and taking into account that all the sub-periods are shorter than the 2006–2017 period, is that the degree of correlation is dependent on the time window analysed. This result echoes those of other authors who have analysed relationships between various oil (crude and product) prices, although for slightly different time intervals [17, 18]. This result could indicate fundamental changes in the energy markets. For example, for the first sub-period (Figure 2b), the correlations are roughly as strong as for the overall period (Figure 2a), though the correlations for propane are stronger than for the overall period. Joint dynamics between crude oil and refined products prices were observed after the onset of the global financial crisis, when oil and refined product prices both decreased in Dec 2008 [17]. In addition to this last result, it can be observed that the correlation values for the second sub-period (Figure 2c) are similar to those of the first (Figure 2b) and to a lesser extent, to those of the overall period (Figure 2a).

For the other two sub-periods (2009–2014, Figure 2d and 2015–2017, Figure 2e) the main feature is the weak correlation (values from 0.30 to 0.43 and from 0.08 to 0.53, respectively) between propane and the other variables, which is most noteworthy in the fourth sub-period (Figure 2e). This can be explained by the following argument, as shown in Figure 1, it can be observed that propane prices have not moved in concordance with those of crude and oil products since approximately 2009. Indeed, there is a “decoupling” of propane prices from the others. Erdős [68] showed that oil and natural gas prices had a long-term equilibrium before 2009 which was decoupled after that. Propane is

obtained mainly from natural gas and its prices are strongly correlated with natural gas prices [60]. On the other hand, [69] suggest that since 2011 WTI prices have ceased to be representative of the price of oil in global markets. Thus, it is clear that propane prices behave in this way. In addition, it is interesting to note that the sub-periods with the lowest correlations values (the third and fourth, Figure 2d and 2e, respectively) coincide with a strong expansion of US stocks (WTI crude oil) and the main driver of this expansion is the increment in tight oil production. This could affect the price dynamics of the energy market system, creating instabilities (volatility, contagion, etc.) in prices. Lastly, some product prices show a relatively low correlation with crude and other product prices, e.g. RBOB gasoline, heating oil, diesel and kerosene. This is more evident in the fourth sub-period (Figure 2e). All of this indicates that the degree of correlation for the period under study is not homogeneous and that the correlation between the oil time series changes over time.

4.4. *Dynamic correlation in the time-frequency domain*

Our WLMC analysis (Figure 3a) shows that the wavelet correlations are strong practically throughout the period under study, particularly the last two wavelet scales (which reach values of almost 1). These wavelet scales are associated with time horizons of between 64 and 128 days and scales (periods) between twice-yearly and two-yearly. This means that they indicate high long-term correlation levels. The correlation coefficients have values ranging from 0.82 (itself a relatively high value) to 0.99 for all the wavelet scales increasing from shorter to longer scales. This result is consistent with the results we obtained previously via the Spearman correlation matrix (see Figure 2a) and with those of other recent publications [18, 25]. However, the WLMC methodology and the results obtained by means of the WLMC bring to light some interesting new points. For instance the WLMC can handle multivariate and multi-time-scale analysis (time and frequency domains) and is able to tackle potentially non-stationary time series [1]. In addition, from a pragmatic point of view, it enables us to analyse the whole period of study under study (2006–2017) without having to split it into several sub-periods.

The most conspicuous and interesting result shown in Figure 3a, is the “strong” decay of the wavelet correlation values (up to 0.82) that takes place approximately from 2013 to 2015 (centred on 2014) for the first three wavelet scales. As discussed earlier in Subsection 4.3, since 2014 WTI crude oil prices have been influenced enormously by non-conventional crude (tight or shale) oil production in the U.S. [70, 71]. Tight oil production has been increasing since 6/2012, but it was not until 2014 that the amount of barrels increased sharply. For this reason, WTI prices and tight oil production show an inverse relationship, i.e. the increase in the tight oil production means a decrease in WTI prices and vice versa. However, there is another potential explanation for this abrupt decline in crude oil prices. A recent study conducted by Kilian [72] finds that the increase in U.S tight oil production is not the main cause of the price decline in U.S. crude oil prices since 2014, which has been driven rather by a combination of positive oil supply shocks, negative shocks to the storage demand for oil (reflecting expectations of lower oil prices) and negative shocks to consumption demand associated with an unexpected slowing of the global economy

[72, 73]. Therefore, a drop in global demand for oil seems to be a major contributor to this crude oil price slump [73, 74].

Our Figure 3a shows that these changes took place at the three shorter wavelet scales associated with time horizons of 1 to 4 days and with daily periods from intraweek to fortnightly. This indicates that the main perturbations in the crude and oil product price system in the short-term are related with “transient” events that appear at most in three weeks. For instance, these drives of “low” correlation could be related to the activities of oil refiners (e.g., adjust of inventory), which produce a break in the connection between crude and product prices [18]. However, clarifying the role of the slowing of the global economy in the oil price system is beyond the scope of our paper. Another reason for the relatively weak short-term correlation could be due a lagged response (e.g., activities of market participants, market power of refineries, unexpected changes in crude oil prices, etc.) of refined product prices [18, 25, 75]. Nevertheless, it is plain to see that there was strong correlation during the global economic depression triggered by the financial (subprime) crisis in 2008. It is widely believed that this crisis resulted in a major crash in energy prices [18]. The three shorter wavelet scales are related to volatility events (e.g., financial process such as contagion) more than to fundamental macroeconomic factors (trade, monetary policy, common shocks, etc.) [39, 76]. On the other hand, it is well known that during financial crisis stock markets tend to be more correlated than in non-crisis periods [38, 39, 77], so it is not unexpected that the oil prices should be strongly correlated during the global economic crisis period.

Some of the most interesting features of WLMC are presented and discussed in the following lines. Our WLMC results (Figure 3b) reveals a very interesting finding: for the first five scales heating oil maximises the multiple correlation of the other oil variables. That is, heating oil can be determined as a linear combination of the rest of the variables for these wavelet scales. This means that, from a theoretical point of view, heating oil is the most dependent variable in the crude/product price system for the first five wavelet scales. This finding is supported by the fact that heating oil prices generally follow crude oil prices [26, 78]. Furthermore, heating oil, kerosene and diesel are the variables with the highest volatility according to their standard deviations (see Table 1). There are several reasons for fluctuations and high volatility in heating oil prices related to time horizons from 1 to 16 days and time periods from intraweek to seasonal scales. For instance, heating oil demand is highly seasonal, which means that it is indirectly influenced by conditions such as severe cold weather events. The mechanism that explains this behaviour is that home heating oil prices tend to rise during the winter when the demand for heating oil is highest. Other factors that affect the oil prices are the competition in local heating oil markets and regional operating costs [78].

By contrast, diesel is the variable that maximises the multiple correlation with the rest of variables in the sixth and seventh scales (associated with time horizons of 32 and 64 days, respectively and time periods between quarterly and two-yearly scales). Like heating oil prices, diesel prices generally follow crude oil prices, but according to the EIA crude oil prices account for about 56% of U.S. diesel fuel prices at the pump from 2007 to 2016. For this reason, crude oil price fluctuations are the main driver of diesel price

fluctuations. However, the wavelet scales indicate that the mechanisms that affect diesel prices occur in the medium and long term. Other mechanisms that can cause diesel price fluctuations are imbalances between U.S. diesel supply and demand (e.g. international demand for diesel affects U.S. diesel prices), seasonality (as for heating oil), transportation costs and local market competition [78].

Finally, kerosene maximises the multiple correlation with the rest of variables at the longest scale (associated with time horizons of 128 days and time periods of 256-512 days, which corresponds to annual and two-yearly scales). It is worth mentioning that this result is particularly noteworthy due to how few papers have analysed the relationship between crude oil and kerosene. On the one hand, our result is in partial agreement with [79], who find that correlations between WTI and kerosene increase with the frequency of data (from weekly to yearly), indicating that there is a co-movement in their price behaviour. Moreover, [26] show that there is a linear trend for WTI and Kerosene, indicating a more stable, more constant trend between crude oil and kerosene. This suggests that there is a long term relationship between these two price time series, so WTI and kerosene prices can be said to be co-integrated [79]. Our WLMC result also indicates that kerosene can be expressed as a linear combination of the rest of the prices for the longest scale. This means that at least WTI and kerosene prices cannot move in opposite directions for long periods without reverting to a long-run equilibrium. This result can also be interpreted and explained in the following manner: on the basis of the results explained above, it can be established that it is possible to hedge markets using WTI futures instead of kerosene futures given that the latter is not liquid. This information is relevant for traders and investors such as kerosene hedgers. For instance, an airline company needs to protect itself against adverse movements in jet fuel prices. Jet fuel futures markets are not liquid, so airlines use crude oil futures to hedge their price risks. The efficiency of this approach is determined by the extent to which the series commove during the time interval of the price exposure.

5. Conclusions

This paper proposes the use of an innovative multivariate and dynamic approach – wavelet local multiple correlation (WLMC) [1] – to analyse the relationship between oil time series in the time-scale domain. This approach is suitable for use with any kind of energy data that change over time (non-stationary) and involve heterogeneous agents who make decisions on different time horizons and operate on different time scales. As a case study, we presents a dynamic analysis of the prices of seven commodities: crude oil (WTI) and six distilled products (conventional gasoline, RBOB regular gasoline, heating oil, Ultra-Low-Sulphur diesel fuel, kerosene and propane) from 10/06/2006 to 17/01/2017. This period contains several financially extreme economic events, including the global financial crisis and increases in tight oil production.

As an exploratory analysis, we estimate a correlation matrix using Spearman’s rank correlation coefficients for the variables under study to test whether the correlations between oil time series change over time. We estimate the correlation for the whole

period of study and for four sub-periods chosen in such a way that they cover certain financial/energy events of interest. We chose the Spearman method because it does not assume normal distribution for the variables analysed and oil time series are not necessarily normally distributed, but other estimators of correlation can be used. The first conspicuous result obtained from the Spearman correlation is that, in general, the correlation between all the variables under analysis is strong (with values from 0.89 to 0.99) with the exception of propane, which shows the weakest correlations with the other variables; the second result is that the highest correlations (between 0.91 and 0.99) are those between products (excluding propane) rather than those with crude oil prices; and finally, the third result is that the degree of correlation between time series is dependent on the time window analysed, indicating that the degree of correlation for the period of study is not homogeneous. That is, the correlation between oil time series changes over time, suggesting the need for a more sophisticated statistical technique for analysing non-stationary data. In addition to the previous conclusions, we highlight that correlation estimated by time intervals is cumbersome in practice.

The wavelet local multiple correlation results reveal that the wavelet correlations are strong, for practically the whole period under study, particularly for the two longest wavelet scales (with correlation values of almost 1). However, the most interesting WLMC result is the strong decay in wavelet correlation values (to values of 0.82) that takes place approximately from 2013 to 2015 (centred on 2014) for the first three wavelet scales. The most plausible explanation for this decay is overproduction of non-conventional crude oil (tight or shale) in the U.S. and a slowdown in global demand for oil. The WLMC also is able to provide a very interesting and original feature, i.e, it is highlighted the variable that maximises the multiple wavelet local correlation in each wavelet scale. Following this characteristic of the WLMC, our WLMC results also reveal that heating oil maximises multiple correlation with other oil variables for the first five scales. By contrast diesel and kerosene are the variables that maximises multiple correlation with the rest of variables at the longest (sixth/seventh and eight) scales. This can be interpreted as meaning that heating oil, diesel or kerosene can be seen as linear combination of the rest of the variables for these wavelet scales. This means that the heating oil, diesel and kerosene are the most dependent variable in the crude/product prices system for the three shorter and the two longest wavelet scales, respectively.

The results obtained using WLMC could be used to draw up a more efficient period-oriented strategy when the market participants have different temporal horizons (investors, speculators, hedge funds, refineries, etc.) and also for energy policy. In addition, the proposed methodology can be used to improve hedging performance depending on the temporal horizon. For instance, a better understanding of the relationships between the prices of crude oil and distilled products is vital for the energy sector, investors, economists and policymakers, especially since the onset of the global economic crisis, particularly as regards the sharp growth in tight oil production in recent times, recognized as one of the most serious extreme economical/energy events ever reported. Thus, these results are relevant to obtaining a better understanding of the behaviour of energy markets.

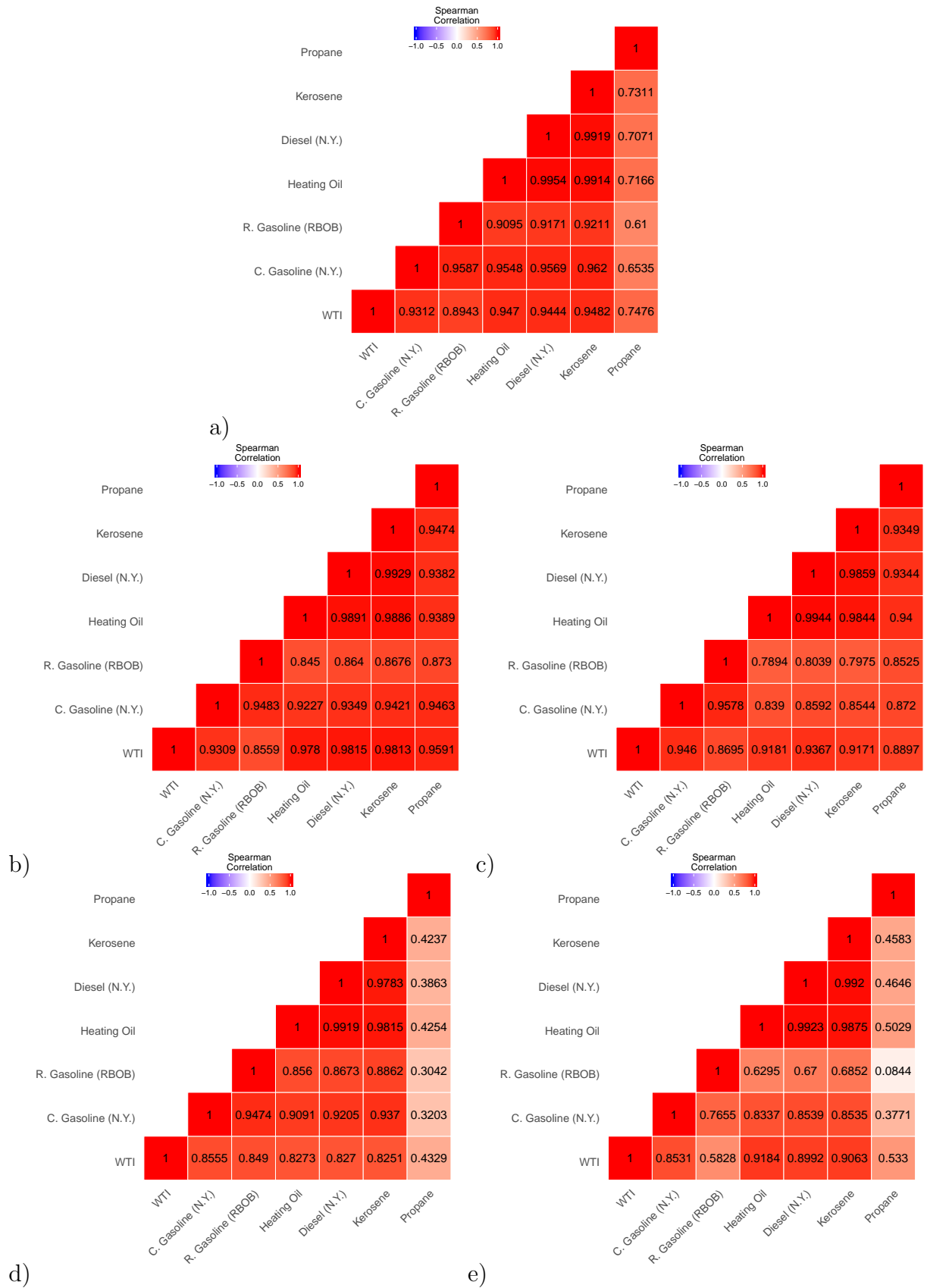
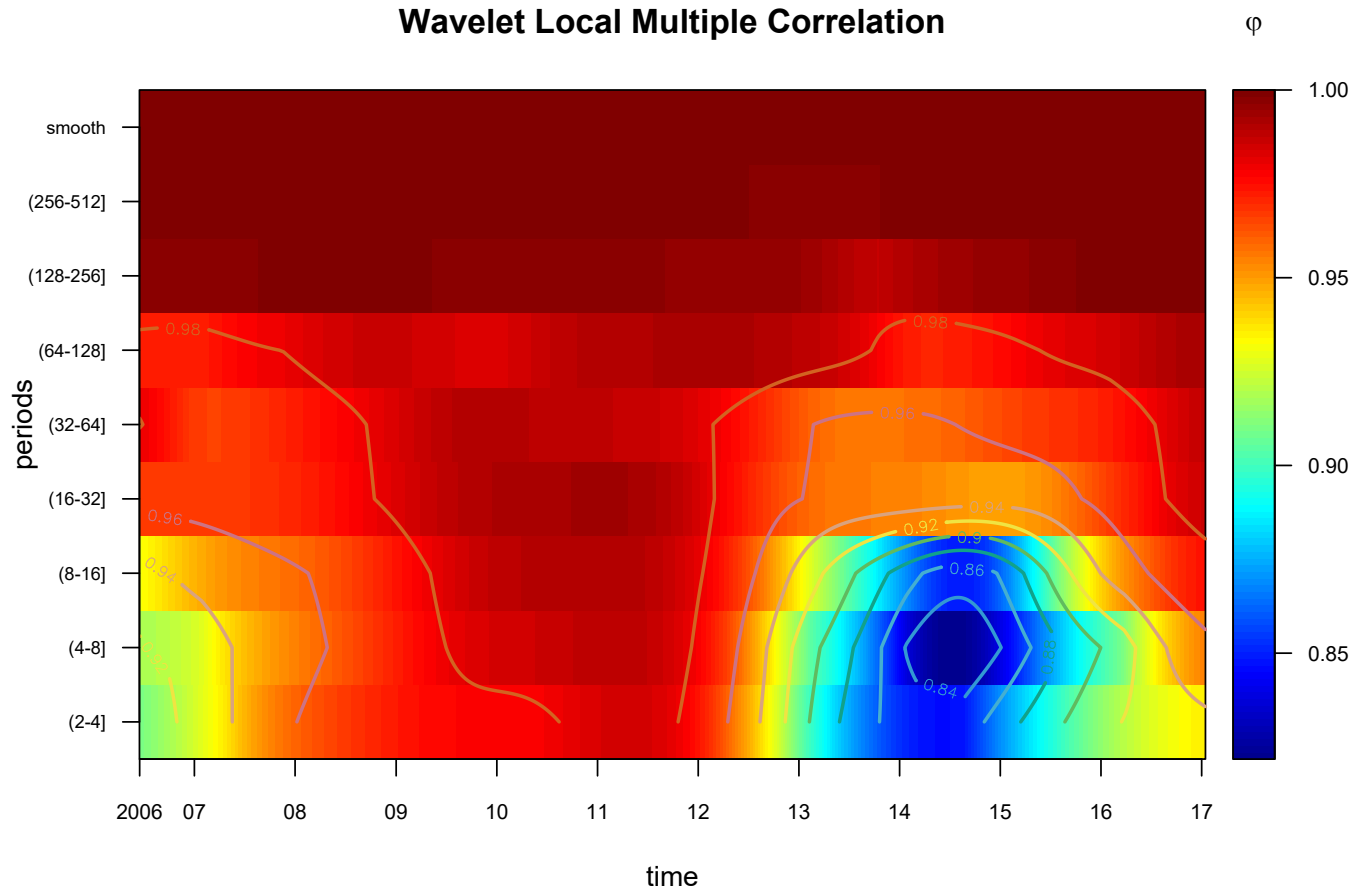


Figure 2: Spearman's correlation matrix for the oil time series under study for the periods: 14/06/2006–17/01/2017 (a), 16/06/2006–31/12/2008 (b), 01/06/2008–31/05/2009 (c), 01/01/2009–31/12/2014 (d) and 01/01/2015–16/02/2017 (e)

a)



b)

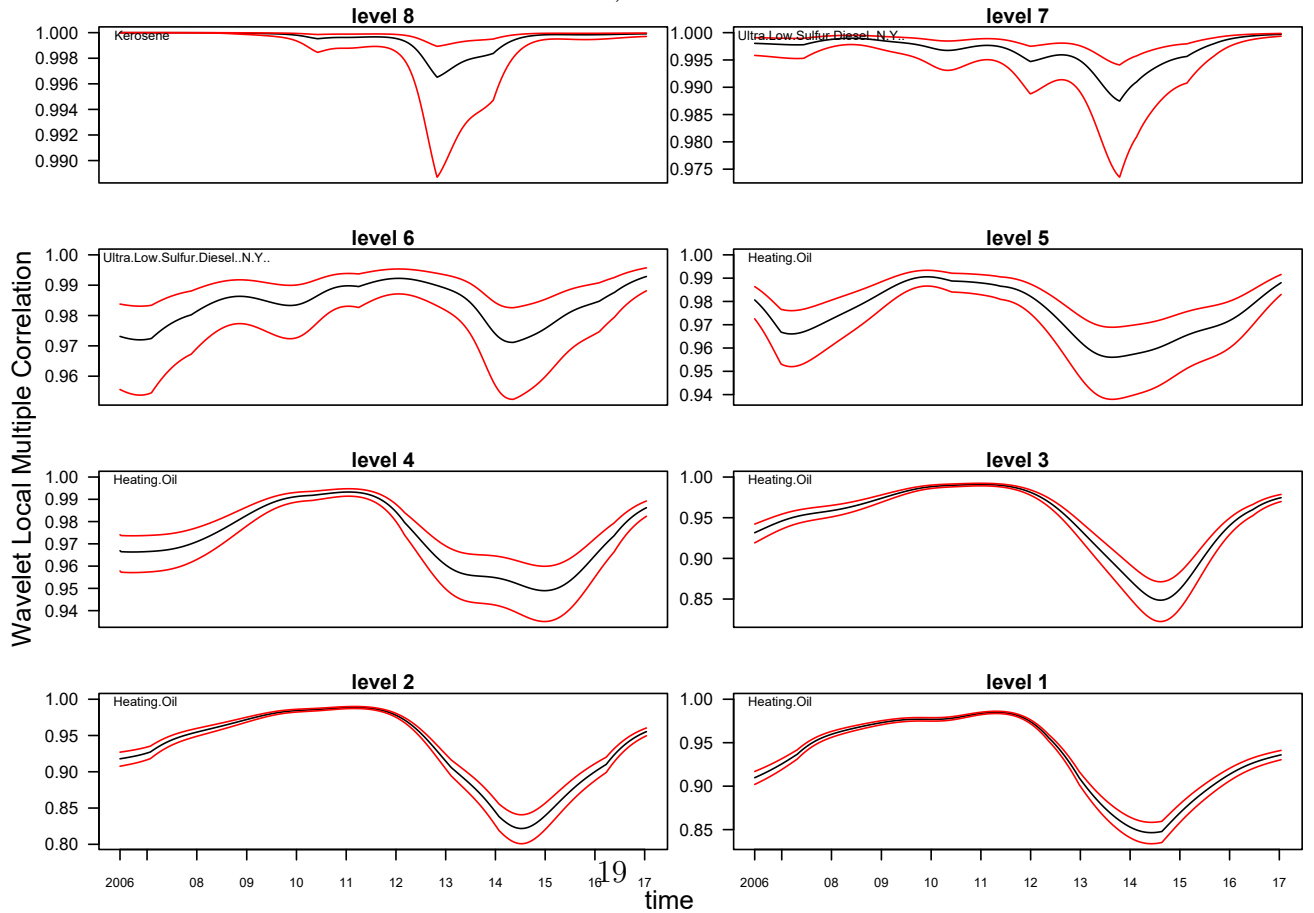


Figure 3: Wavelet local multiple correlation (WMC) for WTI and oil product prices for the period 14/06/2006 – 17/01/2017 (daily data). The red lines show the upper and lower bounds of the 95% confidence interval and the variable named at each upper-left corner gives the highest multiple correlation vs. a linear combination of the rest.

Acknowledgements

JMPM was funded by a Basque Government post-doctoral fellowship. LMA gratefully acknowledges financial support from the Basque Government IT799-13 and from the Spanish Ministry of Science and Innovation (ECO2015-68023). JFM would like to acknowledge research funding received from UPV/EHU Econometrics Research Group (Basque Government Dpt. of Education grant IT-642-13) and Spanish Ministry of Economy and Competitiveness (grant MTM2013-40941-P). We thank to the Editors and the anonymous referees for their valuable comments, which helped us to improve the manuscript. We thanks to BA Christopher Pellow for his English language corrections.

References

- [1] J. Fernández-Macho, Time-localized wavelet multiple regression and correlation, *Physica A: Statistical Mechanics and its Applications* 492 (2018) 1226–1238.
- [2] H. H. Lean, M. McAleer, W.-K. Wong, Market efficiency of oil spot and futures: A mean-variance and stochastic dominance approach, *Energy Economics* 32 (5) (2010) 979–986.
- [3] X. Jia, H. An, X. Sun, X. Huang, L. Wang, Evolution of world crude oil market integration and diversification: A wavelet-based complex network perspective, *Applied Energy* 185 (2017) 1788–1798.
- [4] H. An, X. Gao, W. Fang, Y. Ding, W. Zhong, Research on patterns in the fluctuation of the co-movement between crude oil futures and spot prices: A complex network approach, *Applied Energy* 136 (2014) 1067–1075.
- [5] W. Mensi, S. Hammoudeh, S.-M. Yoon, Structural breaks, dynamic correlations, asymmetric volatility transmission, and hedging strategies for petroleum prices and usd exchange rate, *Energy Economics* 48 (2015) 46–60.
- [6] A. Lanza, M. Manera, M. Giovannini, Modeling and forecasting cointegrated relationships among heavy oil and product prices, *Energy Economics* 27 (6) (2005) 831–848.
- [7] J. Población, G. Serna, Is the refining margin stationary?, *International Review of Economics & Finance* 44 (2016) 169–186.
- [8] A. García-Mirantes, J. Población, G. Serna, Analyzing the dynamics of the refining margin: Implications for valuation and hedging, *Quantitative Finance* 12 (12) (2012) 1839–1855.
- [9] L. Yu, Z. Wang, L. Tang, A decomposition–ensemble model with data-characteristic-driven reconstruction for crude oil price forecasting, *Applied Energy* 156 (2015) 251–267.

- [10] T. Yao, Y.-J. Zhang, C.-Q. Ma, How does investor attention affect international crude oil prices?, *Applied Energy* 205 (2017) 336–344.
- [11] Y.-J. Zhang, Speculative trading and WTI crude oil futures price movement: an empirical analysis, *Applied Energy* 107 (2013) 394–402.
- [12] X. Zhang, L. Yu, S. Wang, K. K. Lai, Estimating the impact of extreme events on crude oil price: An EMD-based event analysis method, *Energy Economics* 31 (5) (2009) 768–778.
- [13] S.-S. Chen, K.-W. Hsu, Reverse globalization: Does high oil price volatility discourage international trade?, *Energy Economics* 34 (5) (2012) 1634–1643.
- [14] K. He, L. Yu, K. K. Lai, Crude oil price analysis and forecasting using wavelet decomposed ensemble model, *Energy* 46 (1) (2012) 564–574.
- [15] Z. A. Ozdemir, K. Gokmenoglu, C. Ekinci, Persistence in crude oil spot and futures prices, *Energy* 59 (2013) 29–37.
- [16] F. Asche, O. Gjølberg, T. Völker, Price relationships in the petroleum market: an analysis of crude oil and refined product prices, *Energy Economics* 25 (3) (2003) 289–301.
- [17] B. Tong, C. Wu, C. Zhou, Modeling the co-movements between crude oil and refined petroleum markets, *Energy Economics* 40 (2013) 882–897.
- [18] L. Liu, G. Ma, Cross-correlation between crude oil and refined product prices, *Physica A: Statistical Mechanics and its Applications* 413 (2014) 284–293.
- [19] M. Cummins, A. Bucca, Quantitative spread trading on crude oil and refined products markets, *Quantitative Finance* 12 (12) (2012) 1857–1875.
- [20] J. P. Byrne, M. Lorusso, B. Xu, Oil prices, fundamentals and expectations, *Energy Economics* doi:<https://doi.org/10.1016/j.eneco.2018.05.011>.
- [21] P. B. Girma, A. S. Paulson, Risk arbitrage opportunities in petroleum futures spreads, *Journal of Futures Markets* 19 (8) (1999) 931–955.
- [22] O. Gjølberg, T. Johnsen, Risk management in the oil industry: can information on long-run equilibrium prices be utilized?, *Energy Economics* 21 (6) (1999) 517–527.
- [23] A. Murat, E. Tokat, Forecasting oil price movements with crack spread futures, *Energy Economics* 31 (1) (2009) 85–90.
- [24] A. Honarvar, Asymmetry in retail gasoline and crude oil price movements in the united states: an application of hidden cointegration technique, *Energy Economics* 31 (3) (2009) 395–402.

- [25] T. Zhang, G. Ma, G. Liu, Nonlinear joint dynamics between prices of crude oil and refined products, *Physica A: Statistical Mechanics and its Applications* 419 (2015) 444–456.
- [26] A. Block Souza, M. B. Righi, S. G. Schlender, D. A. Coronel, Investigating dynamic conditional correlation between crude oil and fuels in non-linear framework: The financial and economic role of structural breaks, *Energy Economics* 49 (2015) 23–32.
- [27] A. A. Suleymanov, A. A. Abbasov, A. J. Ismaylov, Fractal analysis of time series in oil and gas production, *Chaos, Solitons & Fractals* 41 (5) (2009) 2474–2483.
- [28] J. Alvarez-Ramirez, M. Cisneros, C. Ibarra-Valdez, A. Soriano, Multifractal Hurst analysis of crude oil prices, *Physica A: Statistical Mechanics and its Applications* 313 (3) (2002) 651–670.
- [29] E. Panas, V. Ninni, Are oil markets chaotic? A non-linear dynamic analysis, *Energy Economics* 22 (5) (2000) 549–568.
- [30] R. Cont, Empirical properties of asset returns: stylized facts and statistical issues, *Quantitative Finance* 1 (2) (2001) 223–236.
- [31] J. M. Polanco-Martínez, L. Abadie, Analyzing crude oil spot price dynamics versus long term future prices: A wavelet analysis approach, *Energies* 9 (12) (2016) 1089. doi:<http://dx.doi.org/10.3390/en9121089>.
- [32] A. K. Tiwari, C. T. Albulescu, Oil price and exchange rate in india: Fresh evidence from continuous wavelet approach and asymmetric, multi-horizon granger-causality tests, *Applied Energy* 179 (2016) 272–283.
- [33] R. Gençay, F. Selçuk, B. Whitcher, *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics*, Academic Press, 2002.
- [34] D. Percival, A. Walden, *Wavelet methods for time series analysis*, Cambridge University Press, 2000, reprinted 2006.
- [35] M. Gallegati, Wavelet analysis of stock returns and aggregate economic activity, *Computational Statistics & Data Analysis* 52 (6) (2008) 3061–3074.
- [36] J. Fernández-Macho, Wavelet multiple correlation and cross-correlation: A multi-scale analysis of Eurozone stock markets, *Physica A: Statistical Mechanics and its Applications* 391 (4) (2012) 1097–1104.
- [37] J. M. Polanco-Martínez, F. J. Fernández-Macho, Package W2CWM2C: Description, features, and applications, *Computing in Science & Engineering* 16 (6) (2014) 68–78.

- [38] J. Polanco-Martínez, J. Fernández-Macho, M. Neumann, S. Faria, A pre-crisis vs. crisis analysis of peripheral EU stock markets by means of wavelet transform and a nonlinear causality test, *Physica A: Statistical Mechanics and its Applications* 490 (2018) 1211–1227.
- [39] F. Benhmad, Bull or bear markets: A wavelet dynamic correlation perspective, *Economic Modelling* 32 (2013) 576–591.
- [40] J. Fernández-Macho, *wavemulcor*: Wavelet routines for global and local multiple correlation, R package version 2.1.0 (2017).
URL <https://CRAN.R-project.org/package=wavemulcor>
- [41] G. S. Uddin, A. K. Tiwari, M. Arouri, F. Teulon, On the relationship between oil price and exchange rates: A wavelet analysis, *Economic Modelling* 35 (2013) 502–507.
- [42] J. C. Reboredo, M. A. Rivera-Castro, Wavelet-based evidence of the impact of oil prices on stock returns, *International Review of Economics & Finance* 29 (2014) 145–176.
- [43] S. Huang, H. An, X. Gao, X. Sun, Do oil price asymmetric effects on the stock market persist in multiple time horizons?, *Applied Energy* 185 (2017) 1799–1808.
- [44] S. Huang, H. An, X. Huang, X. Jia, Co-movement of coherence between oil prices and the stock market from the joint time-frequency perspective, *Applied Energy* 221 (2018) 122–130.
- [45] S. Huang, H. An, X. Gao, X. Huang, Time–frequency featured co-movement between the stock and prices of crude oil and gold, *Physica A: Statistical Mechanics and its Applications* 444 (2016) 985–995.
- [46] L. Aguiar-Conraria, M. J. Soares, The continuous wavelet transform: moving beyond uni-and bivariate analysis, *Journal of Economic Surveys* 28 (2) (2014) 344–375.
- [47] O. R. Olayeni, Causality in continuous wavelet transform without spectral matrix factorization: theory and application, *Computational Economics* 47 (3) (2016) 321–340.
- [48] V. L. Tonn, H. Li, J. McCarthy, Wavelet domain correlation between the futures prices of natural gas and oil, *The Quarterly Review of Economics and Finance* 50 (4) (2010) 408–414.
- [49] B. P. Van Milligen, E. Sanchez, T. Estrada, C. Hidalgo, B. Brañas, B. Carreras, L. Garcia, Wavelet bicoherence: a new turbulence analysis tool, *Physics of Plasmas* 2 (8) (1995) 3017–3032.

- [50] M. Alzahrani, M. Masih, O. Al-Titi, Linear and non-linear Granger causality between oil spot and futures prices: A wavelet based test, *Journal of International Money and Finance* 48 (2014) 175–201.
- [51] X. Jia, H. An, W. Fang, X. Sun, X. Huang, How do correlations of crude oil prices co-move? a grey correlation-based wavelet perspective, *Energy Economics* 49 (2015) 588–598.
- [52] J.-L. Deng, Control problems of grey systems, *Systems & Control Letters* 1 (5) (1982) 288–294.
- [53] M. Ranta, Contagion among major world markets: a wavelet approach, *International Journal of Managerial Finance* 9 (2) (2013) 133–149. doi:doi:10.1108/17439131311307556.
- [54] A. Tiwari, M. Mutascu, C. Albuлесcu, Continuous wavelet transform and rolling correlation of European stock markets, *International Review of Economics & Finance* 42 (2016) 237–256.
- [55] F. Benhmad, Bull or bear markets: A wavelet dynamic correlation perspective, *Economic Modelling* 32 (2013) 576–591.
- [56] J. Fernández-Macho, Wavelet multiple correlation and cross-correlation: A multi-scale analysis of Eurozone stock markets, *Physica A: Statistical Mechanics and its Applications* 391 (2012) 1097–1104.
- [57] D. Percival, A. Walden, *Wavelet methods for time series analysis*, Cambridge University Press, 2000, reprinted 2006.
- [58] B. Whitcher, P. Gutterp, D. Percival, Wavelet analysis of covariance with application to atmospheric time series, *Journal of Geophysical Research* 105 (2000) 941–962.
- [59] R. Gençay, F. Selçuk, B. Whitcher, *An introduction to wavelets and other filtering methods in finance and economics*, Academic Press, 2002.
- [60] M. Sloan, 2016 propane market outlook: Key market trends, opportunities, and threats facing the consumer propane industry through 2025. Propane Education and Research Council, ICF International, VA, USA. (2016).
- [61] A. García Mirantes, J. Población, G. Serna, Analyzing the dynamics of the refining margin: Implications for valuation and hedging, *Quantitative Finance* 12 (12) (2012) 1839–1855.
- [62] A. Trapletti, K. Hornik, B. LeBaron, Tseries: time series analysis and computational finance, R package version 0.10-11.
- [63] R. N. Mantegna, H. E. Stanley, *Introduction to Econophysics: correlations and Complexity in Finance*, Cambridge university press, 1999.

- [64] R. Zadourian, P. Grassberger, Asymmetry of cross-correlations between intra-day and overnight volatilities, *EPL (Europhysics Letters)* 118 (1) (2017) 18004.
- [65] G. Zebende, M. Da Silva, A. Machado Filho, DCCA cross-correlation coefficient differentiation: Theoretical and practical approaches, *Physica A: Statistical Mechanics and its Applications* 392 (8) (2013) 1756–1761.
- [66] M. Mudelsee, *Climate Time Series Analysis: Classical Statistical and Bootstrap Methods*, Springer, 2014.
- [67] Oil price dot com, <https://oilprice.com/Energy/Heating-Oil/Heating-Oil-This-Essential-Fuel-Explained.html>, accessed: 2017-12-23.
- [68] P. Erdős, Have oil and gas prices got separated?, *Energy Policy* 49 (2012) 707–718.
- [69] C. Baumeister, L. Kilian, Forty years of oil price fluctuations: Why the price of oil may still surprise us, *The Journal of Economic Perspectives* 30 (1) (2016) 139–160.
- [70] R. Alquist, J.-D. Guénette, A blessing in disguise: The implications of high global oil prices for the North American market, *Energy Policy* 64 (2014) 49–57.
- [71] L. Kilian, The impact of the shale oil revolution on us oil and gasoline prices, *Review of Environmental Economics and Policy* 10 (2) (2016) 185–205.
- [72] L. Kilian, The impact of the fracking boom on Arab oil producers, *The Energy Journal* 38 (6) (2017) 137–160.
- [73] L. Kilian, How the tight oil boom has changed oil and gasoline markets, *Papeles de Energía* 3 (2017) 79–102.
- [74] C. Baumeister, L. Kilian, Understanding the decline in the price of oil since June 2014, *Journal of the Association of Environmental and Resource Economists* 3 (1) (2016) 131–158.
- [75] S. Radchenko, D. Shapiro, Anticipated and unanticipated effects of crude oil prices and gasoline inventory changes on gasoline prices, *Energy Economics* 33 (5) (2011) 758–769.
- [76] M. Gallegati, A wavelet-based approach to test for financial market contagion, *Computational Statistics & Data Analysis* 56 (11) (2012) 3491–3497.
- [77] G.-J. Wang, C. Xie, M. Lin, H. E. Stanley, Stock market contagion during the global financial crisis: A multiscale approach, *Finance Research Letters* 22 (2017) 163–168.
- [78] U.S. Energy Information Administration, https://www.eia.gov/energyexplained/index.cfm?page=heating_oil_factors_affecting_prices, accessed: 2018-01-10.

- [79] I. Figuerola-Ferreti, The term structure of cross correlations between the kerosene and crude oil markets, <https://www.rcem.eu/views-on-energy-news/the-term-structure-of-cross-correlations-between-the-kerosene-and-crude-oil-markets> accessed: 2018-01-17.